EPJ manuscript No.
(will be inserted by the editor)

Average path length for Sierpinski pentagon

Junhao Peng[1,2 a] and Guoai Xu[2]

[1] College of Math and Information Science , Guangzhou University , Guangzhou 510006 ,Peoples Republic of China.

[2] State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications , Beijing 100876 , Peoples Republic of China.

the date of receipt and acceptance should be inserted later

Abstract. In this paper,we investigate diameter and average path length(APL) for Sierpinski pentagon based on its recursive construction and self-similar structure.We find that the diameter of Sierpinski pentagon is just the shortest path length between two nodes of generation 0. Deriving and solving the linear homogenous recurrence relation the diameter satisfies,we obtain rigorous solution for the diameter .We also obtain approximate solution for APL of Sierpinski pentagon, both diameter and APL grow approximately as a power-law function of network order $N(t)$,with the exponent equals $\frac{\ln(1+\sqrt{3})}{\ln(5)}$. Although the solution for APL is approximate,it is trusted because we have calculated all items of APL accurately except for the compensation(Δ_t) of total distances between non-adjacent branches($\Lambda_t^{1,3}$) ,which is obtained approximately by least-squares curve fitting.The compensation(Δ_t) is only a small part of total distances between non-adjacent branches($\Lambda_t^{1,3}$) and has little effect on APL.Further more,using the data obtained by iteration to test the fitting results,we find the relative error for Δ_t is less than 10^{-7} ,hence the approximate solution for average path length is almost accurate.

PACS. 02.10.Ox Combinatorics graph theory – 89.75.Hc Networks and genealogical trees – 06.30.Bp Spatial dimensions

1 Introduction

Recently,complex networks have attracted a surge of interest from the scientific community [1–5]. Most endeavors were devoted to unveil the structural properties of real network systems ,such as degree distribution [4–6] ,degree correlation [7, 8],clustering coefficient [9, 10],spectral properties [11–14],diameter [15, 16],average path length [17, 18],communicability [19, 20],etc.These properties play

[a] *Email:* pengjh@gzhu.edu.cn

significant roles in characterizing and understanding complex network systems .

Among these structural properties, average path length (APL) ,which is the mean of the shortest path lengths between all pairs of vertices,characterizes the small-world behavior commonly observed in real networks [17, 18],it is also related to other structural properties, such as degree distribution [21, 22],fractality [23, 24],etc. Further more, average path length has an important consequence for dynamical processes taking placing on networks, including disease spreading [17, 25–27], routing [28–30], percolation [31] , target search [32, 33],and so on .Thus lots of endeavors were devoted to uncovered the APL of different networks,such as Watts-Strogatz model [34], Barabási-Albert network [35],Apollonian network [36],Sierpinski network [37] and hierarchical scale-free network [38],etc.

Sierpinski pentagon belongs to the famous family of Sierpinski objects [39, 40],Lots of job was devoted to study the properties of these objects which also include Sierpinski gasket [41–45],Sierpinski carpet [46–48]and Sierpinski lattice [49, 50] ,etc.As for Sierpinski pentagon, to the best of our knowledge, related research was rarely reported ,and the analytical solution for average path length has not been addressed.

To fill this gap, in this paper, we investigate and obtain approximate solution for average path length . The analytic method is based on the recursive construction and self-similar structure of Sierpinski pentagon. Our results show that the average path length increases approximately algebraically with network order .Although the solution for APL is approximate,it is trusted because we have calculated all items of APL accurately except for the compensation(Δ_t) of total distances between non-adjacent branches($\Lambda_t^{1,3}$).Further more ,the relative errors for Δ_t is less than 10^{-7} as Sec.4.2 shows.

In the process of calculating average path length,we also find that the diameter of Sierpinski pentagon is just the shortest path length between two nodes of generation 0 which has been proved in sec.3.We derive difference equation to depict the evolution of diameter ,solving the difference equation ,we gain the rigorous result which shows that the diameter also increases algebraically with network order .

2 Brief introduction to Sierpinski pentagon

Sierpinski pentagon we considered is a fractal which can be constructed iteratively [39, 40]. We denote the Sierpinski pentagon after t iterations by $G(t)$ with $t \geq 0$.Then the fractal is constructed as follows. For $t = 0, G(0)$ is a filled regular pentagon.In order to obtain $G(1)$,we divide the regular pentagon $G(0)$ so that 6 inner pentagons can be drawn out of it, paint all the inner pentagons but the middle one.Apply the same process to the inner pentagons but the middle one, Sierpinski pentagon is the limiting set for this construction.In this paper,the number of iterations for Sierpinski pentagon is called the generation of Sierpinski pentagon .The Sierpinski pentagon for the first four generation is shown in Figure 1. According to the construction of Sierpinski pentagon, one can see that at each step t, the total number of edges in the systems increases

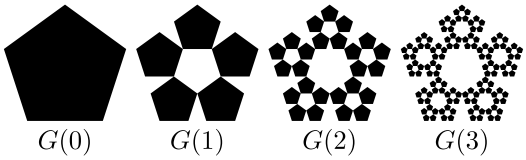

Fig. 1. Growth process for Sierpinski pentagon from generation 0 to generation 3

by a factor of 5. Thus, the total number of edges for $G(t)$ is $E_t = 5^{t+1}$.We can also find that the total number of nodes which is denoted by $N(t)$ satisfies

$$N(t) = 5 \cdot N(t-1) - 5$$

Notice $N(0) = 5$,we can obtain

$$N(t) = \frac{3}{4} \cdot 5^{t+1} + \frac{5}{4} \tag{1}$$

3 Analytical solution of Diameter

The diameter of a graph is the maximum of the shortest path lengths between any pair of nodes,For Sierpinski pentagon,its self-similar structure allows one to find and calculate diameter analytically.The self-similar structure is obvious from an equivalent network construction method: to obtain $G(t+1)$, one can make five copies of $G(t)$ and join them at the five nodes(i.e. , A,B,C,D and E in Figure 2). We can see that the $G(t+1)$ is obtained by the juxtaposition of 5 copies of $G(t)$ which are labeled as $G_1(t)$,$G_2(t)$,$G_3(t)$,$G_4(t)$ and $G_5(t)$,respectively.

We label the five nodes of generation 0 as $1, 2, 3, 4, 5$ and let $d_{ij}(t)$ denotes the shortest path length from node i to j in $G(t)$, it is easy to know: $d_{12}(0) = 1, d_{12}(1) = 4$,$d_{13}(0) = 2, d_{13}(1) = 5$,and for any $t > 1$

$$d_{12}(t) = d_{1A}(t-1) + d_{A2}(t-1) = 2d_{13}(t-1) \tag{2}$$

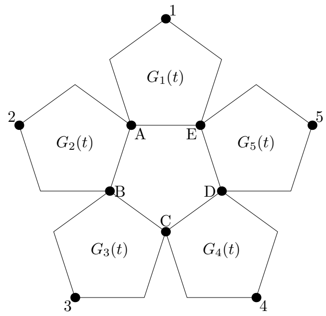

Fig. 2. Second construction method of $G(t)$ that highlights self-similarity. The Sierpinski pentagon $G(t+1)$ is composed of five copies of $G(t)$ denoted as $G_1(t)$,$G_2(t)$,$G_3(t)$,$G_4(t)$,$G_5(t)$.

$$\begin{aligned} d_{13}(t) &= d_{1A}(t-1) + d_{AB}(t-1) + d_{B3}(t-1) \\ &= 2d_{13}(t-1) + d_{12}(t-1) \end{aligned} \tag{3}$$

It follows that

$$d_{13}(t) = 2d_{13}(t-1) + 2d_{13}(t-2) \tag{4}$$

$$d_{12}(t) = 2d_{12}(t-1) + 2d_{12}(t-2) \tag{5}$$

Both $d_{12}(t)$ and $d_{13}(t)$ satisfies the same linear homogenous recurrence relation

$$y_t = 2y_{t-1} + 2y_{t-2}$$

whose general solution [51] is

$$y_t = c_1 \cdot \lambda_1^t + c_2 \cdot \lambda_2^t \tag{6}$$

where λ_1, λ_2 is two roots of its characteristic equation $\lambda^2 - 2\lambda - 2 = 0$,and c_1, c_2 is determined by its initial conditions.Sloving the characteristic equation,we have

$$\lambda_1 = 1 + \sqrt{3}, \lambda_2 = 1 - \sqrt{3}$$

Using the initial conditions $d_{12}(0) = 1, d_{12}(1) = 4$,$d_{13}(0) = 2, d_{13}(1) = 5$,we have

$$d_{12}(t) = \frac{1}{2} \cdot (1+\sqrt{3})^{t+1} + \frac{1}{2} \cdot (1-\sqrt{3})^{t+1} \tag{7}$$

$$d_{13}(t) = \frac{2+\sqrt{3}}{2}\cdot(1+\sqrt{3})^t + \frac{2-\sqrt{3}}{2}\cdot(1-\sqrt{3})^t$$
$$= \frac{1}{4}\cdot(1+\sqrt{3})^{t+2} + \frac{1}{4}\cdot(1-\sqrt{3})^{t+2} \quad (8)$$

In the infinite system size, i.e.,$t \to \infty$

$$d_{13}(t) \approx \frac{(1+\sqrt{3})^{t+2}}{4}$$
$$= \frac{1+\sqrt{3}}{4}\cdot[\frac{4}{3}(N(t)-\frac{5}{4})]^{\frac{\ln(1+\sqrt{3})}{\ln(5)}}$$
$$\propto N(t)^{\frac{\ln(1+\sqrt{3})}{\ln(5)}} \quad (9)$$

As the diameter of a graph is the maximum of the shortest path lengths between any pair of its nodes,we find that the diameter of Sierpinski pentagon $G(t)$ is just $d_{13}(t)$ which has prooved in Theorem 1.Thus, the diameter grows approximately as a power-law function of network order $N(t)$, with the exponent is $\frac{\ln(1+\sqrt{3})}{\ln(5)}$.

Theorem 1 *For Sierpinski pentagon,let* $L(t)$ *denote the diameter of Sierpinski pentagon* $G(t)$*,thus,* $L(t) = d_{13}(t)$.

Proof:In fact,we want to proof that the inequality

$$d_{ij}(t) \le d_{13}(t) \quad (10)$$

holds for any $t > 0$,and any two nodes i, j in $G(t)$.

Here we prove the result by mathematical induction .

Initial step:For t=0,it is easy to know that the inequality (10) holds.

Inductive step:Assume there is a $k \ge 0$, such that inequality (10) holds for $t = k$, we must prove the inequality (10) holds for $t = k + 1$.

For any two nodes i, j of $G(k+1)$,if i, j belong to the same $G(k+1)$ branch which is a copy of $G(k)$,Thus inequality (10) holds because

$$d_{ij}(k+1) = d_{ij}(k) \le d_{13}(k) < d_{13}(k+1)$$

If i, j belong to two different branches of $G(k+1)$,it can be discussed on two cases according the relation of the two different branches.

I)If the two branches is adjacent,by symmetry,we can suppose that i belongs to $G_1(t)$,j belongs to $G_2(t)$,the inequality (10) holds because

$$d_{ij}(k+1) = d_{iA}(k+1) + d_{Aj}(k+1)$$
$$\le d_{1A}(k+1) + d_{A2}(k+1)$$
$$< d_{13}(k+1)$$

II)If the two branches is not adjacent,by symmetry,we can suppose that i belongs to $G_1(t)$,j belongs to $G_3(t)$,we have

$$d_{ij}(k+1) \le d_{iA}(k+1) + d_{AB}(k+1) + d_{Aj}(k+1)$$
$$\le d_{1A}(k+1) + d_{AB}(k+1) + d_{A2}(k+1)$$
$$= d_{13}(k+1)$$

Thus inequality (10) holds for this case which finish the proof.

4 Derivation of Average path length

We represent all the shortest path lengths of the Sierpinski pentagon $G(t)$ as a matrix in which the entry $d_{ij}(t)$ is the shortest distance from node i to node j, and d_t denotes the average path length (APL) of $G(t)$ which is defined as the mean of $d_{ij}(t)$ over all couples of nodes,thus

$$d_t = \frac{D_t}{N(t)(N(t)-1)/2} \quad (11)$$

where

$$D_t = \sum_{i,j\in G(t), i\neq j} d_{ij}(t) \quad (12)$$

denotes the sum of the shortest path length between two nodes over all pairs.

Based on the self-similar structure of $G(t+1)$ As shown in Figure.2, it is easy to see that the total distance D_t satisfies

$$D_{t+1} = 5D_t + \Lambda_t \tag{13}$$

where Λ_t ,named the crossing distance in this paper ,denotes the sum over all shortest paths whose end points are not in the same branch.

Let $\Lambda_t^{i,j}$ denotes the sum of all shortest paths whose endpoints are in $G_i(t)$ and $G_j(t)$ excluding paths whose end nodes are in the same branch,and $D_t^{i,\alpha}$ denotes the sum of all shortest paths from hub node α (i.e., A,B,C,D and E in figure2)to any nodes in $G_i(t)$,that is to say

$$\Lambda_t^{1,2} = \sum_{\substack{i\in G_1(t), j\in G_2(t)\\ i,j\neq A}} d_{ij}(t) \tag{14}$$

$$\Lambda_t^{1,3} = \sum_{\substack{i\in G_1(t), j\in G_3(t)\\ i\neq A, j\neq B}} d_{ij}(t) \tag{15}$$

$$D_t^{1,C} = \sum_{i\in G_1(t)} d_{iC}(t) \tag{16}$$

It is easy to know from the self-similar structure of $G(t+1)$

$$\begin{aligned}\Lambda_t = \quad & \Lambda_t^{1,2} + \Lambda_t^{2,3} + \Lambda_t^{3,4} + \Lambda_t^{4,5} + \Lambda_t^{5,1} \\ & + \Lambda_t^{1,3} + \Lambda_t^{1,4} + \Lambda_t^{2,4} + \Lambda_t^{2,5} + \Lambda_t^{3,5} \\ & - D_t^{1,C} - D_t^{2,D} - D_t^{3,E} - D_t^{4,A} - D_t^{5,B}\end{aligned} \tag{17}$$

The last five terms of Eq. (17) which want to be subtracted are the items which have been calculated twice .For example,$D_t^{1,C}$ (excluding $d_{AC}(t)$)is calculated both in $\Lambda_t^{1,3}$ and $\Lambda_t^{1,4}$,and $d_{AC}(t)$ is calculated both in $\Lambda_t^{2,3}$ and $\Lambda_t^{1,4}$. By symmetry,we have

$$\Lambda_t = 5\Lambda_t^{1,2} + 5\Lambda_t^{1,3} - 5D_t^{1,C} \tag{18}$$

4.1 Total distances from one node of generation 0 to any other nodes

In this section ,we will calculate a quantity which will be used in calculating Λ_t ,the quantity denoted by S_t is the total distances from one node of generation 0 (labeled by $1,2,3,4,5$ which was shown in Figure.2)to any other nodes of $G(t)$ .It is easy to know that, this quantity is equal for any nodes of generation 0,thus

$$S_t = \sum_{i\in G(t), i\neq 5} d_{i5}(t) \tag{19}$$

It is easy to know that $S_0 = 6$. We also find that S_t satisfies the recurrence relations derived as follows,which will help us to obtain the analytical solution for S_t .

Note that $G(t+1)$ is obtained by the juxtaposition of 5 copies of $G(t)$ which are labeled as $G_1(t)$,$G_2(t)$,$G_3(t)$,$G_4(t)$ and $G_5(t)$,respectively,we have

$$\begin{aligned}S_{t+1} = & \sum_{i\in G_5(t), i\neq 5} d_{i5}(t+1) + \sum_{i\in G_1(t), i\neq E} d_{i5}(t+1) \\ & + \sum_{i\in G_4(t), i\neq D} d_{i5}(t+1) + \sum_{i\in G_2(t), i\neq A} d_{i5}(t+1) \\ & + \sum_{i\in G_3(t), i\neq C} d_{i5}(t+1) - d_{B5}(t+1)\end{aligned} \tag{20}$$

where

$$\sum_{i\in G_5(t), i\neq 5} d_{i5}(t+1) = S_t$$

$$\begin{aligned}\sum_{i\in G_1(t), i\neq E} d_{i5}(t+1) &= \sum_{i\in G_1(t), i\neq E} (d_{iE}(t+1) + d_{E5}(t+1)) \\ &= S_t + (N(t)-1)d_{13}(t)\end{aligned}$$

$$\sum_{i\in G_4(t), i\neq D} d_{i5}(t+1) = S_t + (N(t)-1)d_{13}(t)$$

$$\sum_{i\in G_2(t), i\neq A} d_{i5}(t+1) = S_t + (N(t)-1)(d_{12}(t) + d_{13}(t))$$

$$\sum_{i\in G_3(t), i\neq C} d_{i5}(t+1) = S_t + (N(t)-1)(d_{12}(t) + d_{13}(t))$$

$$d_{B5}(t) = d_{BA}(t) + d_{AE}(t) + d_{E5}(t) = 2d_{12}(t) + d_{13}(t)$$

Thus

$$\begin{aligned} S_t &= 5S_{t-1} + (2N(t-1)-4)d_{12}(t-1) \\ &+ (4N(t-1)-5)d_{13}(t-1) \\ &\equiv 5S_{t-1} + f(t-1) \\ &= 5(5S_{t-2} + f(t-2)) + f(t-1) \\ &= 5^2 S_{t-2} + 5f(t-2) + f(t-1) \\ &= \dots \\ &= 5^t S_0 + 5^{t-1} f(0) + \dots + 5f(t-2) + f(t-1) \\ &= 5^t \cdot 6 + \sum_{i=0}^{t-1}[5^{t-1-i} f(i)] \end{aligned} \tag{21}$$

with

$$f(t) \equiv (2N(t)-4)d_{12}(t) + (4N(t)-5)d_{13}(t)$$

It follows from Eqs.(1),(7) and (8) that

$$\begin{aligned} &\sum_{i=0}^{t-1}[5^{t-1-i} f(i)] \\ &= \sum_{i=0}^{t-1} 5^{t-1-i}[(2N(i)-4)d_{12}(i) + (4N(i)-5)d_{13}(i)] \\ &= \sum_{i=0}^{t-1} 5^{t-1-i}[(\frac{3}{2}5^{i+1} - \frac{3}{2})d_{12}(i) + 3\cdot 5^{i+1} d_{13}(i)] \\ &= \frac{3}{2}\sum_{i=0}^{t-1} 5^t d_{12}(i) - \frac{3}{2}\sum_{i=0}^{t-1} 5^{t-1-i} d_{12}(i) + 3\sum_{i=0}^{t-1} 5^t d_{13}(i) \\ &= \frac{3}{2}\cdot 5^t \sum_{i=0}^{t-1}[\frac{1}{2}\cdot(1+\sqrt{3})^{i+1} + \frac{1}{2}\cdot(1-\sqrt{3})^{i+1}] \\ &\quad -\frac{3}{2}\sum_{i=0}^{t-1} 5^{t-1-i}[\frac{1}{2}\cdot(1+\sqrt{3})^{i+1} + \frac{1}{2}\cdot(1-\sqrt{3})^{i+1}] \\ &\quad +3\cdot 5^t \sum_{i=0}^{t-1}[\frac{1}{4}\cdot(1+\sqrt{3})^{i+2} + \frac{1}{4}\cdot(1-\sqrt{3})^{i+2}] \\ &= \frac{3}{2}5^t[\frac{1+\sqrt{3}}{2}\frac{(1+\sqrt{3})^t - 1}{\sqrt{3}} + \frac{1-\sqrt{3}}{2}\frac{(1-\sqrt{3})^t - 1}{-\sqrt{3}}] \\ &\quad -\frac{3}{2}5^{t-1}[\frac{1+\sqrt{3}}{2}\frac{(\frac{1+\sqrt{3}}{5})^t - 1}{\frac{1+\sqrt{3}}{5} - 1} + \frac{1-\sqrt{3}}{2}\frac{(\frac{1-\sqrt{3}}{5})^t - 1}{\frac{1-\sqrt{3}}{5} - 1}] \\ &\quad +3\cdot 5^t[\frac{2+\sqrt{3}}{2}\frac{(1+\sqrt{3})^t - 1}{\sqrt{3}} + \frac{2-\sqrt{3}}{2}\frac{(1-\sqrt{3})^t - 1}{-\sqrt{3}}] \\ &= 5^t[\frac{9+5\sqrt{3}}{4}(1+\sqrt{3})^t + \frac{9-5\sqrt{3}}{4}(1-\sqrt{3})^t - \frac{69}{13}] \\ &\quad +\frac{21+15\sqrt{3}}{52}(1+\sqrt{3})^t + \frac{21-15\sqrt{3}}{52}(1-\sqrt{3})^t \end{aligned} \tag{22}$$

Hence

$$\begin{aligned} S_t &= 5^t[\frac{9+5\sqrt{3}}{4}(1+\sqrt{3})^t + \frac{9-5\sqrt{3}}{4}(1-\sqrt{3})^t + \frac{9}{13}] \\ &+ \frac{21+15\sqrt{3}}{52}(1+\sqrt{3})^t + \frac{21-15\sqrt{3}}{52}(1-\sqrt{3})^t \end{aligned} \tag{23}$$

4.2 Total distances of Non-adjacent branch: $\Lambda_t^{1,3}$

Now,we will derive the total distances between branch $G_1(t)$ and $G_3(t)$ which is denoted by $\Lambda_t^{1,3}$. According to the construction shown in Figure.2,we can find that, for most pairs of nodes $i,j(i \in G_1(t), j \in G_3(t))$, the shortest path pass through node A and B ,hence

$$\begin{aligned} \Lambda_t^{1,3} &= \sum_{\substack{i\in G_1(t), j\in G_3(t)\\ i\neq A, j\neq B}} d_{ij}(t) \\ &= \sum_{\substack{i\in G_1(t), j\in G_3(t)\\ i\neq A, j\neq B}} [d_{iA}(t) + d_{AB}(t) + d_{Bj}(t)] - \Delta_t \\ &= \sum_{\substack{i\in G_1(t), j\in G_3(t)\\ i\neq A, j\neq B}} d_{iA}(t) + \sum_{\substack{i\in G_1(t), j\in G_3(t)\\ i\neq A, j\neq B}} d_{AB}(t) \\ &\quad + \sum_{\substack{i\in G_1(t), j\in G_3(t)\\ i\neq A, j\neq B}} d_{Bj}(t) - \Delta_t \\ &= (N(t)-1)\sum_{i\in G_1(t), i\neq A} d_{iA}(t) + (N(t)-1)^2 d_{AB}(t) \\ &\quad +(N(t)-1)\sum_{j\in G_2(t), j\neq A} d_{Bj}(t) - \Delta_t \\ &= 2(N(t)-1)S_t + (N(t)-1)^2 d_{12}(t) - \Delta_t \end{aligned} \tag{24}$$

with Δ_t to compensate for the overcount of certain pairs whose shortest paths does not pass through A, B.while t shows that both branch $G_1(t)$ and $G_3(t)$ are a copy of Sierpinski pentagon $G(t)$.

In order to calculating Δ_t,we must know when overcount occurs and how many it is.For any pair of nodes

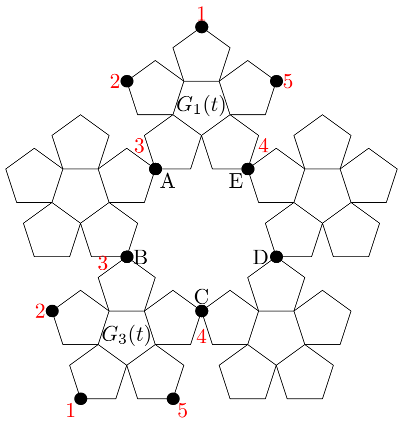

Fig. 3. The branch $G_1(t)$ and $G_3(t)$ which is looked upon as a Sierpinski pentagon $G(t)$ with the five nodes of generation 0 labeled as $1, 2, 3, 4$ and 5.

$i, j (i \in G_1(t), j \in G_3(t))$, the shortest path can pass through node A and B or pass through C, D and E,let $d^A_{ij}(t), d^E_{ij}(t)$ denote the two kinds of shortest path between i, j in $G(t)$, respectively , we have

$$\begin{aligned} d^A_{ij}(t+1) &= d_{iA}(t+1) + d_{AB}(t+1) + d_{Bj}(t+1) \\ &= d_{iA}(t+1) + d_{Bj}(t+1) + d_{12}(t) \\ d^E_{ij}(t+1) &= d_{iE}(t+1) + d_{ED}(t+1) \\ &\quad + d_{DC}(t+1) + d_{Cj}(t+1) \\ &= d_{iE}(t+1) + d_{Cj}(t+1) + 2d_{12}(t) \end{aligned}$$

Thus

$$\begin{aligned} d^A_{ij}(t+1) - d^E_{ij}(t+1) &= d_{iA}(t+1) - d_{iE}(t+1) \\ &\quad + d_{Bj}(t+1) - d_{Cj}(t+1) - d_{12}(t) \end{aligned}$$

If $d^A_{ij}(t+1) - d^E_{ij}(t+1) > 0$, overcount occurs ,and it will be added into Δ_t,if $d^A_{ij}(t+1) - d^E_{ij}(t+1) \leq 0$,it has no effect on Δ_t.

If we look upon branch $G_1(t)$ and $G_3(t)$ as a Sierpinski pentagon $G(t)$ and label the five nodes of generation 0 as $1, 2, 3, 4$ and 5, which is shown in Figure 3, we find that the hub node A ,E ,B,C of $G(t+1)$ is just node $3, 4$ of $G_1(t)$ and node $3, 4$ of $G_3(t)$,while $G_1(t)$ and $G_3(t)$ is looked upon as a Sierpinski pentagon $G(t)$ .Thus

$$\begin{aligned} &d^A_{ij}(t+1) - d^E_{ij}(t+1) \\ &= d_{i3}(t) - d_{i4}(t) + d_{j3}(t) - d_{j4}(t) - d_{12}(t) \end{aligned} \tag{25}$$

which imply that $d^A_{ij}(t+1) - d^E_{ij}(t+1)$ subjects to $d_{i3}(t) - d_{i4}(t)$,while $d_{i3}(t) - d_{i4}(t)$ is the distance difference for node i to node 3 and 4 in Sierpinski pentagon $G(t)$. We can calculate Δ_t If we can obtain $d_{i3}(t) - d_{i4}(t)$ for all nodes i of $G(t)$,which can be solved based on the recurrence relations $d_{i3}(t) - d_{i4}(t)$ satisfies.

According the construction of $G(t+1)$.we can find the distance difference from i to node $2, 3, 4$ satisfies recurrence relations which rely on the branch where node i locates.If i is a node of $G_1(t)$

$$d_{i3}(t+1) - d_{i4}(t+1) = d_{i3}(t) - d_{i4}(t) \tag{26}$$

$$d_{i2}(t+1) - d_{i4}(t+1) = d_{i3}(t) - d_{i4}(t) - d_{12}(t) \tag{27}$$

If if i is a node of $G_2(t)$

$$d_{i3}(t+1) - d_{i4}(t+1) = -d_{12}(t) \tag{28}$$

$$\begin{aligned} &d_{i2}(t+1) - d_{i4}(t+1) \\ &= d_{i2}(t) - d_{i4}(t) - d_{12}(t) - d_{13}(t) \end{aligned} \tag{29}$$

If i is a node of $G_3(t)$

$$\begin{aligned} &d_{i3}(t+1) - d_{i4}(t+1) \\ &= d_{i2}(t) - d_{i4}(t) - d_{13}(t) \end{aligned} \tag{30}$$

$$\begin{aligned} &d_{i2}(t+1) - d_{i4}(t+1) \\ &= d_{i2}(t) - d_{i4}(t) - d_{13}(t) - d_{12}(t) \end{aligned} \tag{31}$$

If i is a node of $G_4(t)$

$$d_{i3}(t+1) - d_{i4}(t+1)$$
$$= d_{i2}(t) - d_{i4}(t) + d_{13}(t) \quad (32)$$
$$d_{i2}(t+1) - d_{i4}(t+1)$$
$$= d_{i2}(t) - d_{i4}(t) + d_{13}(t) + d_{12}(t) \quad (33)$$

If i is a node of $G_5(t)$

$$d_{i3}(t+1) - d_{i4}(t+1) = d_{12}(t) \quad (34)$$
$$d_{i2}(t+1) - d_{i4}(t+1)$$
$$= d_{i2}(t) - d_{i3}(t) + d_{12}(t) \quad (35)$$

Let $\Omega_{3,4}(t), \Omega_{2,4}(t)$ denote the set of all the values of $d_{i3}(t) - d_{i4}(t)$ and $d_{i2}(t) - d_{i4}(t)$, respectively. While t=0,$\Omega_{3,4}(0) = \{-1, -1, 0, 1, 1\}$,$\Omega_{2,4}(0) = \{-2, -1, 0, 1, 2\}$. For $t > 0$,we can obtain $\Omega_{3,4}(t), \Omega_{2,4}(t)$ based on the recurrence relations as Eqs.(26)-(35) show.

Now we come back to analyze Δ_t which is the sum of all non-negative values of Eq.(25). It is easy to know that $d_{i3}(t)-d_{i4}(t)$ and $d_{j3}(t)-d_{j4}(t)$ in Eq.(25) has the same set of possible values : $\Omega_{3,4}(t)$ which has just been obtained . We can calculate all the possible values of Eq.(25),and Δ_t is get by adding all the non-negative values,the results for $t = 0 \sim 11$ is shown in Table 1.

But with the increasing of t ,it is difficult to calculate Δ_t by iteration because it is prohibitively time and memory consuming.Substituting Eqs.(1),(7) and (23) into Eq.(24),we find that the expression for $\Lambda_t^{1,3}$ satisfies Eq.(36) if we ignore Δ_t .We also believe Δ_t can only change $\Lambda_t^{1,3}$ a little ,and Δ_t can be approximated by Eq.(36) .

$$\Phi(t) = a_1 \cdot 5^{2t}(1+\sqrt{3})^t + a_2 \cdot 5^{2t}(1-\sqrt{3})^t$$
$$+a_3 \cdot 5^t(1+\sqrt{3})^t + a_4 \cdot 5^t(1-\sqrt{3})^t + a_5 \cdot (1+\sqrt{3})^t$$
$$+a_6 \cdot (1-\sqrt{3})^t + a_7 \cdot 5^{2t} + a_8 \cdot 5^t + a_9 \quad (36)$$

Table 1. The value of Δ_t for $t = 0 \sim 11$ obtained by iteration

t	Δ_t	t	Δ_t	t	Δ_t
0	4	4	3697330	8	79817184975658
1	30	5	251032868	9	$5.45159641 \times 10^{15}$
2	1002	6	171140501308	10	$3.72349326 \times 10^{17}$
3	56540	7	1168705606692	11	$2.54319349 \times 10^{19}$

with the 9 coefficients determined by the actual data shown in Table 1.Using standard software package of MATLAB R2008a ,we obtain the 9 coefficients of Δ_t by least-squares curve fitting .Results show the residual is equal to 1.78×10^{16} and relative error(defined as the absolute error divided by the true value) is also large .If we delete the term 5^{2t} in Δ_t whose fitting coefficient ($a_7 = -0.0011$) is small and conduct least-squares curve fitting again based on data for $t = 0 \sim 7$ in Table 1,the residual is equal to 2.15×10^{-7} which is very small and the 8 coefficients of Δ_t is:

$$a_1 = 0.168524328052979, a_2 = -0.0396624946437528$$
$$a_3 = 0.935344610079585, a_4 = 0.951717329999713$$
$$a_5 = -4.00947432595951, a_6 = -1.49385494978489$$
$$a_8 = 2.71082171547275, a_9 = 4.7765837996533 \quad (37)$$

We compare the results calculated by fitting curve and the data in Table 1 for $t = 0 \sim 11$, the relative error is less than 10^{-7} which is acceptable and can not be avoided for round-off error,thus our model is trusted.The reason why

we don't conduct least-squares curve fitting with more data is the relative error become lager if we use more data for there is huge differences among the data for different t.

4.3 Approximate solution for average path length

In this subsection, we calculate $\Lambda_t^{1,2}$ and $D_t^{1,C}$ first , and then Λ_t and D_t can be obtained from Eqs.(18) and (13).Finally,we calculate average path length from Eq.(11).

According to the construction of $G(t+1)$,we find

$$\begin{aligned}\Lambda_t^{1,2} &= \sum_{\substack{i\in G_1(t),j\in G_2(t)\\ i,j\neq A}} d_{ij}(t+1)\\ &= \sum_{\substack{i\in G_1(t),j\in G_2(t)\\ i,j\neq A}} [d_{iA}(t+1)+d_{Aj}(t+1)]\\ &= 2(N(t)-1)S_t \end{aligned} \tag{38}$$

and

$$\begin{aligned}D_t^{1,C} &= \sum_{i\in G_1(t)} d_{iC}(t+1)\\ &= \sum_{i\in G_1(t)} [2d_{12}(t)+min\{d_{iA}(t+1),d_{iE}(t+1)\}]\\ &= 2N(t)d_{12}(t)+\sum_{i\in G(t)} min\{d_{i3}(t),d_{i4}(t)\}\\ &\equiv 2N(t)d_{12}(t)+F(t) \end{aligned} \tag{39}$$

where

$$\begin{aligned}F(t) &= \sum_{i\in G(t)} min\{d_{i3}(t),d_{i4}(t)\}\\ &= 2S_{t-1}+2[N(t-1)-1]d_{13}(t-1)\\ &\quad +2S_{t-1}-d_{13}(t-1)+F(t-1)\\ &\quad +[N(t-1)-2][d_{13}(t-1)+d_{12}(t-1)]\\ &= 4S_{t-1}+[3N(t-1)-5]d_{13}(t-1)\\ &\quad +[N(t-1)-2]d_{12}(t-1)+F(t-1)\\ &= 4S_{t-1}+[3N(t-1)-5]d_{13}(t-1)\\ &\quad +[N(t-1)-2]d_{12}(t-1)\\ &\quad +4S_{t-2}+[3N(t-2)-5]d_{13}(t-2)\\ &\quad +[N(t-2)-2]d_{12}(t-2)+F(t-2)\\ &= \cdots\\ &= 4\sum_{k=0}^{t-1}S_k+\sum_{k=0}^{t-1}\{[3N(k)-5]d_{13}(k)\}\\ &\quad +\sum_{k=0}^{t-1}\{[N(k)-2]d_{12}(k)\}+F(0)\\ &= 4\sum_{k=0}^{t-1}S_k+\sum_{k=0}^{t-1}\{[3N(k)-5]d_{13}(k)\}\\ &\quad +\sum_{k=0}^{t-1}\{[N(k)-2]d_{12}(k)\}+4\\ &= \frac{485\sqrt{3}+792}{472}5^t(1+\sqrt{3})^t+\frac{792-485\sqrt{3}}{472}\\ &\quad \times 5^t(1-\sqrt{3})^t+\frac{48-\sqrt{3}}{312}(1+\sqrt{3})^t\\ &\quad +\frac{48+\sqrt{3}}{312}(1-\sqrt{3})^t+\frac{9}{13}5^t-\frac{21}{59} \end{aligned} \tag{40}$$

Thus,the crossing distance

$$\begin{aligned}\Lambda_t &= 5\Lambda_t^{1,2}+5\Lambda_t^{1,3}-5D_t^{1,C}\\ &= 5\cdot 2(N(t)-1)S_t+5\cdot[2(N(t)-1)S_t+(N(t)\\ &\quad -1)^2d_{12}(t)-\Delta_t]-5\cdot[2N(t)d_{12}(t)+F(t)]\\ &= 20(N(t)-1)S_t+5(N(t)-1)^2d_{12}(t)\\ &\quad -10N(t)d_{12}(t)-5F(t)-5\Delta_t \end{aligned} \tag{41}$$

Substituting Eqs.(40),(1),(7),(8),(23),(37) into Eq.(41),we obtain

$$\begin{aligned}\Lambda_t &= 426.3358\cdot 5^{2t}(1+\sqrt{3})^t-19.1676\cdot 5^{2t}(1-\sqrt{3})^t\\ &\quad +29.4512\cdot 5^t(1+\sqrt{3})^t-0.7143\cdot 5^t(1-\sqrt{3})^t\\ &\quad +7.1748\cdot(1+\sqrt{3})^t+10.6543\cdot(1-\sqrt{3})^t\\ &\quad +51.9230\cdot 5^{2t}-13.5541\cdot 5^t-22.1032 \end{aligned}$$

$$\equiv \sum_{k=1}^{9} c_k q_k^t \tag{42}$$

with the last line is an abbreviation,and c_k, q_k correspond to appropriate expressions shown above. Thus,the total distance

$$\begin{aligned} D_t &= 5D_{t-1} + \Lambda_{t-1} \\ &= 5^2 D_{t-2} + 5\Lambda_{t-2} + \Lambda_{t-1} \\ &= \cdots \\ &= 5^t D_0 + \sum_{i=0}^{t-1} [5^{t-1-i}\Lambda_i] \\ &= 5^t D_0 + \sum_{i=0}^{t-1} [5^{t-1-i} \sum_{k=1}^{9} c_k q_k^i \\ &= 5^t D_0 + 5^{t-1} \sum_{k=1}^{9}\sum_{i=0}^{t-1} c_k (\frac{q_k}{5})^i \\ &= 5^t D_0 + 5^{t-1} \sum_{k=1}^{9} c_k \frac{1-(\frac{q_k}{5})^t}{1-(\frac{q_k}{5})} \end{aligned} \tag{43}$$

It is easy to know that $D_0 = 15$,substituting c_k, q_k $(k = 1, 2, \cdots, 9)$ for the appropriate expressions in Eq.(42) ,we have

$$\begin{aligned} D_t = &\frac{c_1}{5(4+5\sqrt{3})} \cdot 5^{2t}(1+\sqrt{3})^t \\ &+\frac{c_2}{5(4-5\sqrt{3})} \cdot 5^{2t}(1-\sqrt{3})^t \\ &+\frac{c_3}{5\sqrt{3}} \cdot 5^t(1+\sqrt{3})^t - \frac{c_4}{5\sqrt{3}} \cdot 5^t(1-\sqrt{3})^t \\ &-\frac{c_5}{4-\sqrt{3}} \cdot (1+\sqrt{3})^t - \frac{c_6}{4+\sqrt{3}} \cdot (1-\sqrt{3})^t \\ &+\frac{c_7}{20} \cdot 5^{2t} + [15 - \frac{c_1}{5(4+5\sqrt{3})} - \frac{c_2}{5(4-5\sqrt{3})} \\ &-\frac{c_3}{5\sqrt{3}} + \frac{c_4}{5\sqrt{3}} + \frac{c_5}{4-\sqrt{3}} + \frac{c_6}{4+\sqrt{3}} - \frac{c_7}{20} \\ &+\frac{c_8 t}{5} + \frac{c_9}{4}] \cdot 5^t - \frac{c_9}{4} \\ = &\,6.7350 \cdot 5^{2t}(1+\sqrt{3})^t + 0.8226 \cdot 5^{2t}(1-\sqrt{3})^t \\ &+3.4007 \cdot 5^t(1+\sqrt{3})^t + 0.08248 \cdot 5^t(1-\sqrt{3})^t \\ &-3.1636 \cdot (1+\sqrt{3})^t - 1.8587 \cdot (1-\sqrt{3})^t \\ &+2.5961 \cdot 5^{2t} + (0.859 - 2.71t)5^t + 5.526 \end{aligned} \tag{44}$$

It follows from Eqs.(1) and (11) that

$$\begin{aligned} d_t &= \frac{D_t}{\frac{225}{32}5^{2t} + \frac{45}{16}5^t + \frac{5}{32}} \\ &= \frac{\frac{D_t}{5^{2t}}}{\frac{225}{32} + \frac{45}{16}5^{-t} + \frac{5}{32}5^{-2t}} \end{aligned}$$

In the infinite system size, i.e.,$t \to \infty$

$$\begin{aligned} d_t &\approx 0.9579 \cdot (1+\sqrt{3})^t + 0.1170 \cdot (1-\sqrt{3})^t \\ &\quad +0.4836(\frac{1+\sqrt{3}}{5})^t + 0.0117(\frac{1-\sqrt{3}}{5})^t + 0.3692 \\ &\approx 0.9579 \cdot (1+\sqrt{3})^t \\ &= \frac{0.9579}{1+\sqrt{3}} \cdot [\frac{4}{3}(N(t) - \frac{5}{4})]^{\frac{\ln(1+\sqrt{3})}{\ln(5)}} \\ &\propto N(t)^{\frac{\ln(1+\sqrt{3})}{\ln(5)}} \end{aligned} \tag{45}$$

which implies that APL grows approximately as a power-law function of network order $N(t)$, with the exponent is $\frac{\ln(1+\sqrt{3})}{\ln(5)}$.In contrast to many recently studied network models mimicking real-life systems in nature and society [15–18], Sierpinski pentagons are not small worlds .

5 Conclusion

In this paper,we have obtained rigorously solution for the diameter and approximate solution for average path length ,both diameter and APL of Sierpinski pentagons grow approximately as a power-law function of network order $N(t)$.Although the solution for APL is approximate,it is trusted because we have calculated all items of APL accurately except for the compensation(Δ_t) of total distances between non-adjacent branches($\Lambda_t^{1,3}$),which is obtained approximately by least-square curve fitting.The compensation(Δ_t) is only a small part of total distances between non-adjacent branches($\Lambda_t^{1,3}$) and has little effect on

APL. Further more ,we use the data obtained by iteration to test our fitting results and find the relative error for Δ_t is less than 10^{-7} which is acceptable.Hence the approximate solution for average path length is almost accurate.

Acknowledgment

This research was supported by the National High Technology Research and Development Program("863"Program) of China under Grant No. 2009AA01Z439.

References

1. R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
2. S.N. Dorogvtsev and J.F.F. Mendes , Adv. Phys. **51** ,1079(2002).
3. M.E.J. Newman SIAM Rev. **45** 167(2003).
4. A.-L. Barabási, R. Albert , Science **286** ,509(1999).
5. A.-L. Barabási, R. Albert, H. Jeong, and G. Bianconi , Science **287**, 2115(2000).
6. Z. N. Oltvai, A.-L. Barabsi, Science **298**, 763(2002).
7. M.E.J.Newman, Phys. Rev. Lett. **89** , 208701(2002).
8. P. L.Krapivsky, and S. Redner, Phys. Rev. E **63**, 066123(2001).
9. M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys.Rev. E **64**, 026118(2001).
10. Y.C.Zhang, Z.Z.Zhang, J.H.Guan, and S.G.Zhou., J. Stat. Mech. **03**, P03013(2010).
11. I. Farkas, I. Derenyi, A-L. Barabsi, T. Vicsek, Phys.Rev. E **64**, 026704(2001).
12. Z.Z.Zhang , Y.Qi, S.G.Zhou, Y.Lin, and J.H.Guan, Phys.Rev. E **80**, 016104(2009).
13. S.Jalan , G.M.Zhu , and B.W.Li , Phys. Rev. E **84**, 046107(2011) .
14. A.Yamamoto, S.Yamada, M.Okumura, and M.Machida, Phys. Rev. A **84**, 043642(2011).
15. R. Albert,H. Jeong,and A.-L. Barabási , Nature(London) **401**, 130(1999).
16. A. Barratand, M. Weigt , Eur. Phys. J. B **13**, 547(2000).
17. D.J. Watts, H. Strogatz , Nature (London) **393**, 440(1998).
18. M. Barthélémy and L. A. N. Amaral , Phys. Rev. Lett. **82**, 5180(1999).
19. E. Estrada, N. Hatano, Phys. Rev. E **77**, 036111 (2008).
20. E. Estrada, N. Hatano,Appl. Math. Comp. **214** 500-511(2009).
21. F. Chung and L.Lu, Proc. Nat. Acad. Sci. **99** , 15879(2002).
22. R. Cohen and S. Havlin, Phys. Rev. Lett. **90**, 058701(2003).
23. C. Song , S. Havlin and H. A. Makse, Nat. Phys. **2** ,275(2006).
24. Z. Z. Zhang , S. G. Zhou and T. Zou, Eur. Phys. J. B **56**, 259(2007).
25. R. Pastor-Satorras and A. Vespignani , Phys. Rev. Lett. **86**, 3200(2001).
26. L. Ancel Meyers, M. E. J. Newman, M. Martin, and S. Schrag , Emerging Infectious Diseases **9**, 204(2001).
27. A. L. Lloyd and R. M. May , Science **292**, 1316(2001).
28. G. Yan , T. Zhou, B. Hu, Z. Q. Fu and B. H.Wang , Phys. Rev. E **73** ,046108(2006).

29. Z. Z. Zhang , F. Comellas, G. Fertin , A. Raspaud, L. L. Rong and S. G. Zhou , J. Phys. A: Math. Theor. **41**, 035004(2008).
30. F. Comellas, A. Miralles , J. Phys. A: Math. Theor. **42** , 425001(2009).
31. Z.Z.Zhang, S. G.Zhou, T.Zou and G. S.Chen, J. Stat. Mech. ,P09008(2008).
32. F. Jasch and A. Blumen , Phys. Rev. E **63**, 041108(2001).
33. M. F. Shlesinger , Nature(London) **443**, 281(2006).
34. M. E. J .Newman, and D. J. Watts, Phys. Lett. A **263**, 341(1999a).
35. B.Bollobás and O. Riordan, Combinatorica **24**, 5C34(2004).
36. Z.Z.Zhang ,L.C.Chen ,S. G.Zhou, L.j.Fang, J.h.Guan, T.Zou,, Phys. Rev. E **77** , 017102(2008).
37. Z.Z.Zhang ,L.C.Chen ,S. G.Zhou, L.j.Fang, S.G.Zhou,Y.C.Zhang,J.h.Guan, J. Stat. Mech. **02**, P02034(2009).
38. Z.Z.Zhang ,L. Yuan,S.Y.Gao,S.G.Zhou,J.h.Guan, J. Stat. Mech. , P10022(2009).
39. B.Mandlebrot , *The Fractal Geometry of Nature* (San Francisco: Freeman)(1982).
40. R. Engelking, Wiadomosci matematyczne **26(1)**, 18-24(1984).
41. L.J. Bentz, J. W. Turner, and J. J. Kozak, Phys. Rev. E **82**, 011137(2010).
42. S. H. Liu and A. J. Liu, Phys. Rev. B **32**, 4753 (1995).
43. Y.O. Hayase and T.Ohta, Phys. Rev. Lett. **81**, 1726 (1998).
44. J.J.Kozak and V. Balakrishnan , Phys. Rev. E **65**, 021105(2002).
45. S.G.Ri, H.J.Ruan , Journal of Mathematical Analysis and Applications **380(1)**, 313-322(2011).
46. M. Fritsche, H. E. Roman, and M. Porto, Phys. Rev. E **76**, 061101(2007).
47. P.Y. Hsiao and P. Monceau, Phys. Rev. B **65** ,184427(2002).
48. M. A. Bab, G. Fabricius, and E. V. Albano, Phys. Rev. E **71**, 036139(2005).
49. Y. Liu, Z. Hou, P. M. Hui, and W. Sritrakool, Phys. Rev. B **60** , 13444(1999).
50. A. Ordemann, M. Porto, and H. Ed. Roman, Phys. Rev. E **65**, 021107(2002).
51. Richard.A.Brualdi , *Introductory Combinatorics* ,(Third Edition), (Pearson Education,Inc,1999).